

Supplement

Depth variation and stereo processing tasks in natural scenes

Arvind V Iyer¹ & Johannes Burge^{1,2,3}

¹. Department of Psychology, University of Pennsylvania

². Neuroscience Graduate Group, University of Pennsylvania

³. Bioengineering Graduate Group, University of Pennsylvania

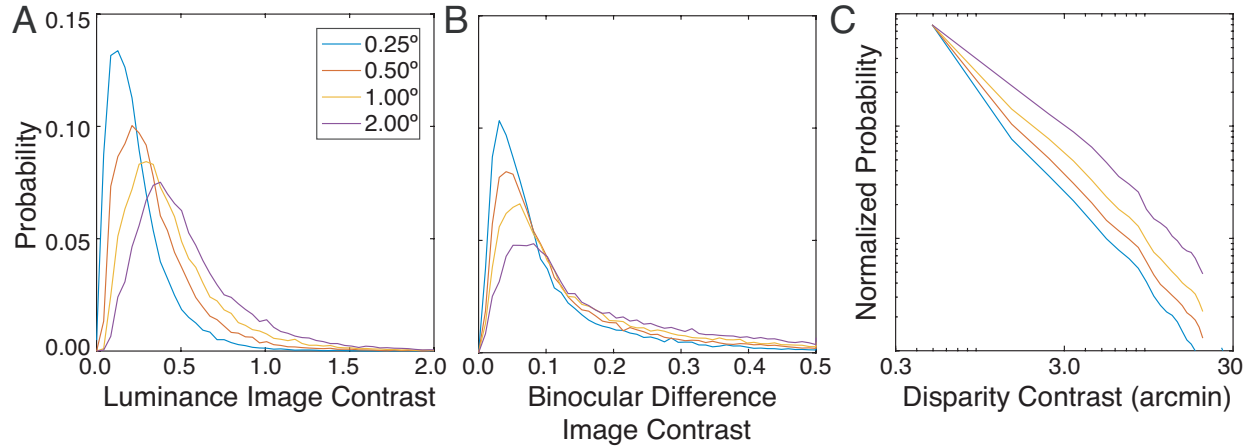


Figure S1. Luminance contrast, binocular difference image contrast, and disparity contrast in natural scenes. **A** Distribution of luminance contrast for different spatial integration areas (colors). Larger luminance contrasts are more likely when computed with larger spatial integration areas (see Methods). **B** Distribution of binocular difference image contrast. Binocular difference image contrast is the contrast of the point-wise difference between the right- and left-eye windowed contrast images (see Methods). Low binocular difference image contrasts are associated with low disparity contrasts (see Fig. 10). **C** Distribution of disparity contrasts. Disparity contrast is approximately distributed as a power law, although the approximation breaks down at very small and large contrasts. The y-axis spans three orders of magnitude. Note that because the database of natural scenes from which these statistics are calculated contain no objects closer than 3.0m, and because the disparity associated with a given depth decreases with the square of distance, these distributions probably underestimate the frequency of high disparity contrasts.

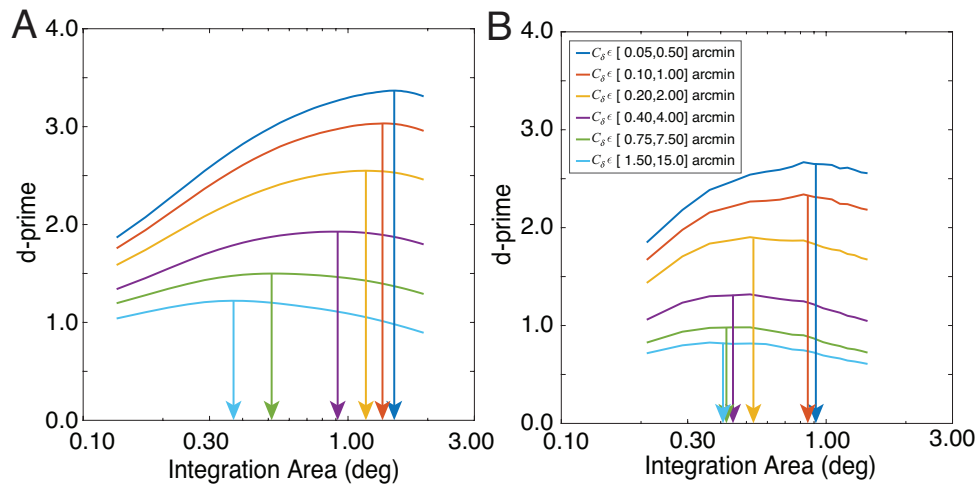


Figure S2: Effect of disparity contrast on task-performance and size of task-optimal integration region **A** Half-occlusion sensitivity as a function of integration area for different disparity contrast intervals (colors). Arrows mark the spatial integration area (window full-width at half-height) for which half-occlusion detection performance is optimized. The analyses are the same as in Fig 5E but for a different choice of disparity contrast bins in arcmin. The dashed curve shows the d-prime obtained for the superset of stimuli in all disparity contrast intervals. **B** Disparity-detection sensitivity as a function of integration area for different disparity contrast intervals (colors). Similar to **A**, but for the disparity detection task (see also Fig 7E)

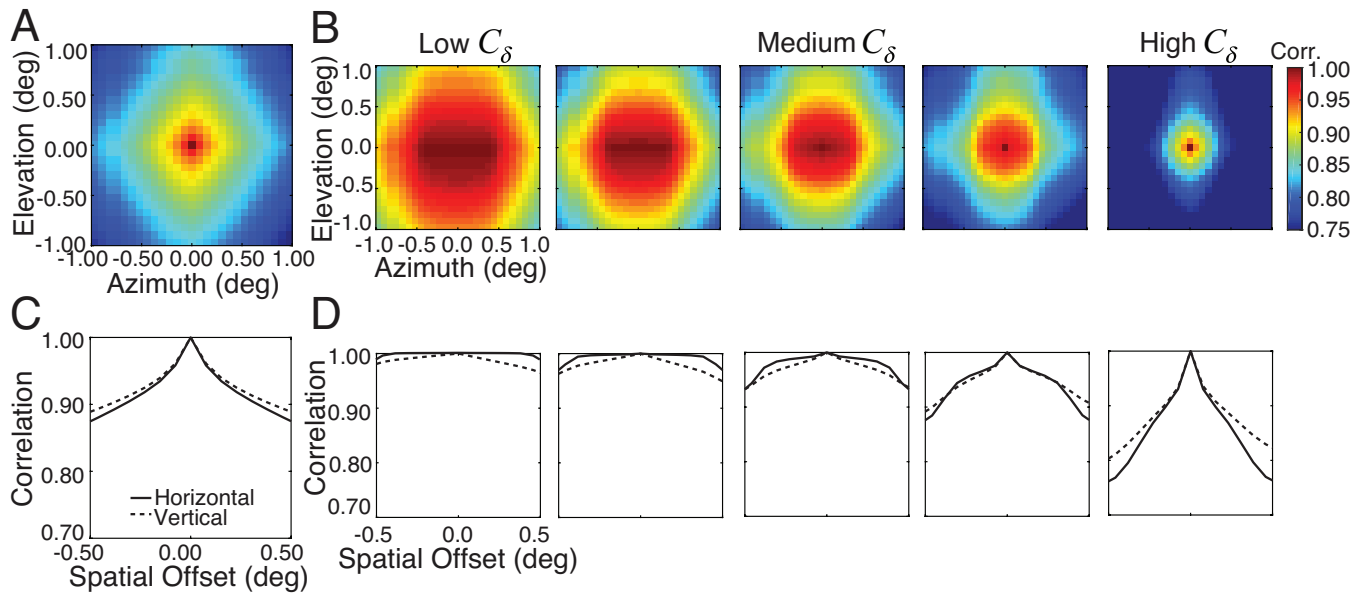


Figure S3. Pixel-based disparity correlation maps. **A** Disparity correlations as a function of retinal position. The value at a location \mathbf{x} is the Pearson correlation coefficient between the vector of disparity values occurring at pixel-location \mathbf{x} across all stimuli, and the vector of all disparity values occurring at the center pixel \mathbf{x}_0 . **B** Same as A, but conditioned on five different disparity contrast bins—0.1-1.0, 0.2-2.0, 0.4-4.0, 0.75-7.5, 1.5-15.0 arcmin—computed within a 1 deg area. At lower disparity contrasts, the correlation remains close to the maximum value of 1.0. . At higher disparity contrasts, the correlation decreases rapidly with eccentricity, and more rapidly with horizontal than with vertical offsets. **C,D** Horizontal and vertical slices through plots in A and B for spatial offsets within 0.5 deg (solid & dashed curves, respectively).

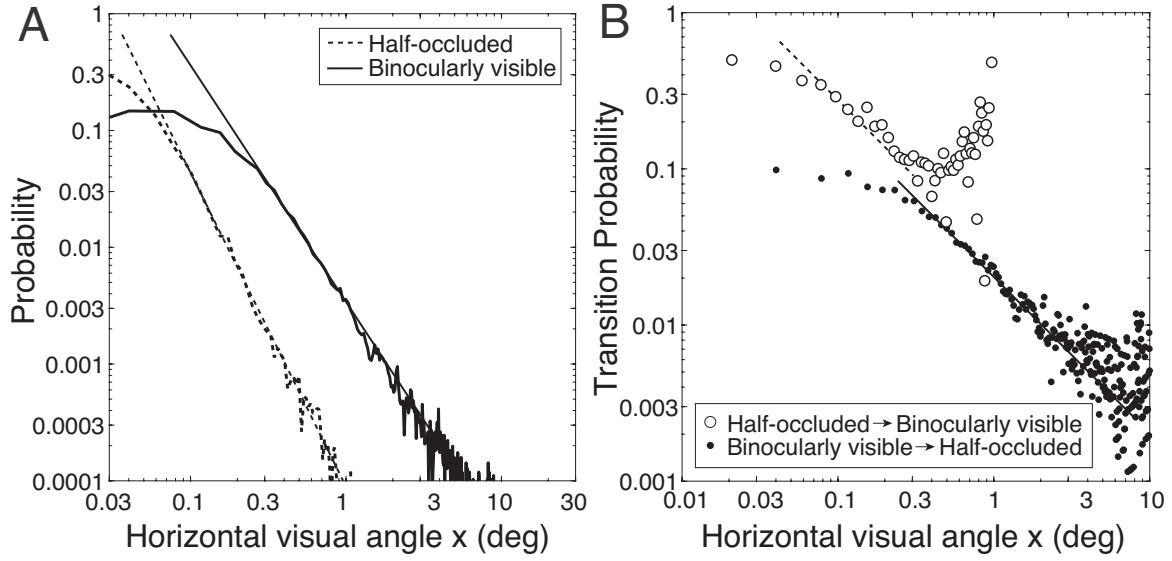


Figure S4. Statistics of binocularly visible and half-occluded regions in natural scenes. **A** Distributions of half-occluded and binocularly visible region sizes computed from all images in the database. The distribution half-occluded region sizes is well described by a power law $p(x_j) = \alpha_j \beta_j^{\alpha_j} / x_j^{\alpha_j+1}$ beyond ~ 0.1 deg (dashed line). The distribution of binocularly visible region sizes is well described by a power law beyond ~ 0.3 deg (solid line). Solid and dashed black lines show the best-fitting power-law fits for the distributions of half-occluded and binocularly visible regions ($\alpha = -2.7$ and $\alpha = -2.0$, respectively). **B** Transition probabilities from half-occluded to binocularly visible regions (white dots) and from binocularly visible to half-occluded regions (black dots) as a function of region size.

Consider a horizontal row of pixels $\mathbf{y} = [y_1, y_2, \dots, y_M]$ where $y_i \in \{0, 1\}$ and 0 denotes a half-occluded pixel and 1 denotes a binocularly visible pixel. Let the random variable X_j denote the length of a contiguous series of pixels of having value $j \in \{0, 1\}$. Specifically, X_0 represents the length of a contiguous string of half-occluded pixels, and X_1 represents the length of a contiguous string of binocularly visible pixels. In general, if X_j has probability density $p_j(x_j)$ and cumulative density $P_j(x_j)$, then the transition probability (i.e. the probability that the next pixel has the opposite label) for each region size is given by $\lambda_j(x_j) = p_j(x_j) / (1 - P_j(x_j))$. If the distribution of contiguous region sizes follows a power-law (e.g. a Pareto distribution of the first kind), then $p(x_j) = \alpha_j \beta_j^{\alpha_j} / x_j^{\alpha_j+1}$ for all $x_j \geq \beta_j$ with shape parameter α_j and scale parameter β_j which represents the minimum size. With this probability density, the transition probabilities are given by the power law $\lambda_j(x_j) = \alpha_j x_j^{-1}$. Thus, the transition probability as a function of region size x_j falls linearly with slope -1.0 on a log-log scale within the region of support (Fig. S4B). This prediction holds for the transition probabilities in our dataset.

For purposes of comparison, if the sequences of half-occluded and binocularly visible points were generated by independent coin flips (i.e. Bernoulli trials) instead of being drawn from the natural scene data, then the region sizes would be exponentially distributed $p_j(x_j) \sim \lambda_j e^{-\lambda_j x_j}$ and the transition probabilities λ_j would be a constant independent of region size. The Pareto distribution provides a better description of the empirical region-size and transition probability data than the exponential distribution (Fig. S4A).

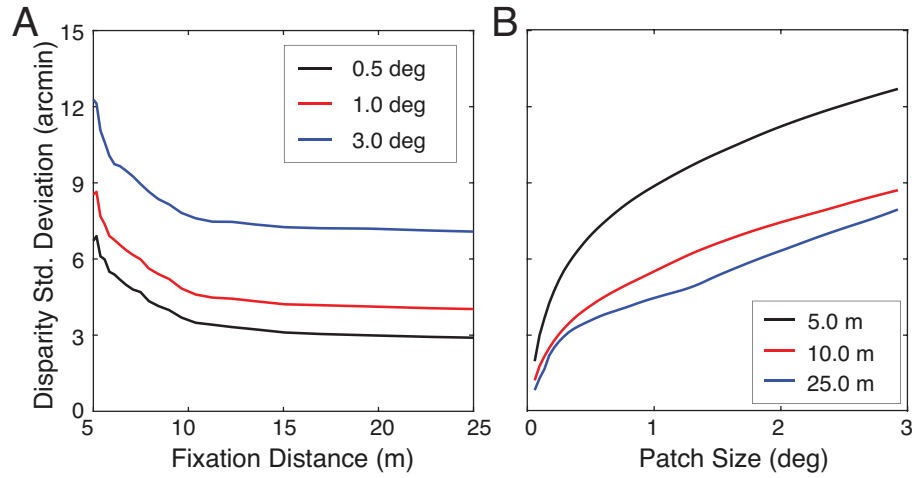


Figure S5: Effect of fixation distance and integration region size on binocular disparity standard deviation. **A** Effect of fixation distance on standard deviation of disparity signals within integration areas of different sizes (colors) centered on the fovea. Disparity standard deviation decreases monotonically with fixation distance. **B** Effect of integration area on disparity standard deviation for three different fixation distances (colors). Disparity standard deviation increases monotonically with integration area. Distance bins were 0.2 diopters wide.

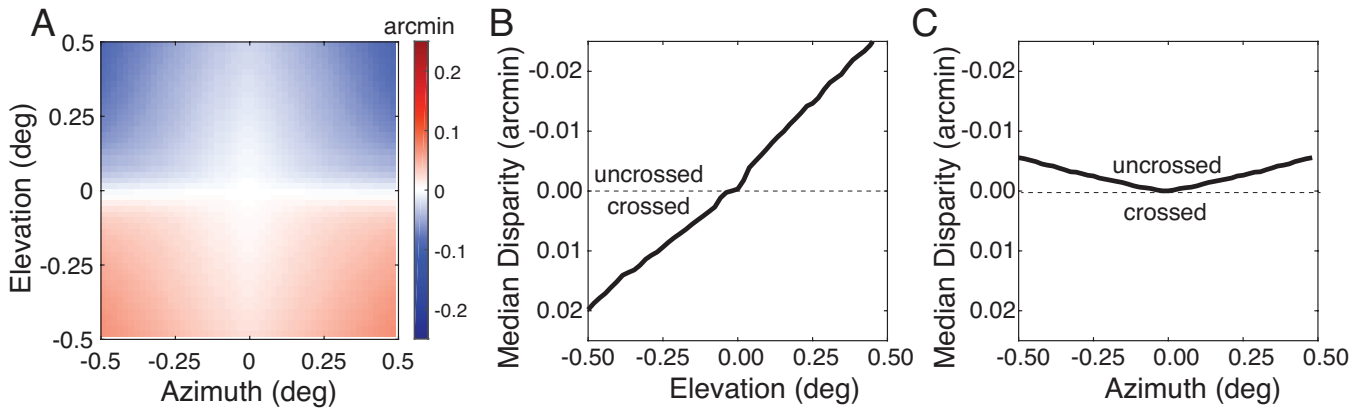


Figure S6. Median disparities in the near foveal region. **A** The spatial map of natural disparity medians is symmetric about the vertical meridian. Uncrossed disparities occur exclusively at elevations above the horizontal median. Crossed disparities occur exclusively at elevations below the horizontal median. This pattern of disparities can be understood in terms of the typical structure of natural scenes where nearer fixations (yielding crossed disparities) occur most frequently at lower elevations where the ground-plane dominates. **B** Vertical horopter. A vertical slice in A at an azimuth of 0.0° shows a backward pitch with respect to the vertical, consistent with a previously reported property of the empirical vertical horopter known as the Helmholtz shear. The best-fit line is $\delta_{median} = -0.0475e - 0.0028$ where e is retinal elevation. The top-back pitch of the vertical horopter is consistent with multiple psychophysical observations. **C** Horizontal horopter. Median disparities along the horizontal median are more uncrossed at larger azimuths. A horizontal slice in A at an elevation of 0.0° is well fit by the parabola $\delta_{median} = -0.022a^2$ where a is retinal azimuth. The horizontal horopter's deviation from the Vieth Müller circle towards uncrossed disparity is consistent with the Hering-Hillebrand deviation.

It has been hypothesized that retinal corresponding points are positioned to be most often stimulated by surfaces in the world (Cooper, Burge, & Banks, 2011; Gibaldi, Canessa, & Sabatini, 2017; Helmholtz, 1925; Schreiber, Hill, Filippini, Schor, & Banks, 2008; Sprague, Cooper, Tonic, & Banks, 2015). This hypothesis accounts for why the empirical vertical horopter is pitched back (i.e. the Helmholtz shear) and why the empirical horizontal horopter deviates from the Vieth-Müller circle (i.e. the Hering-Hillebrand deviation). Under this hypothesis, median disparities along the vertical and horizontal retinal meridians constitute predictions of the empirical horizontal and vertical horopters. We computed the median disparities along the horizontal and vertical retinal meridians in our dataset (Fig. S4A-C). The median disparities exhibit the Helmholtz shear and the Hering-Hillebrand deviation, but smaller in magnitude from psychophysically reported values, consistent with the expectation that disparity reliability is underestimated because of the absence of objects nearer than 3m in our dataset.

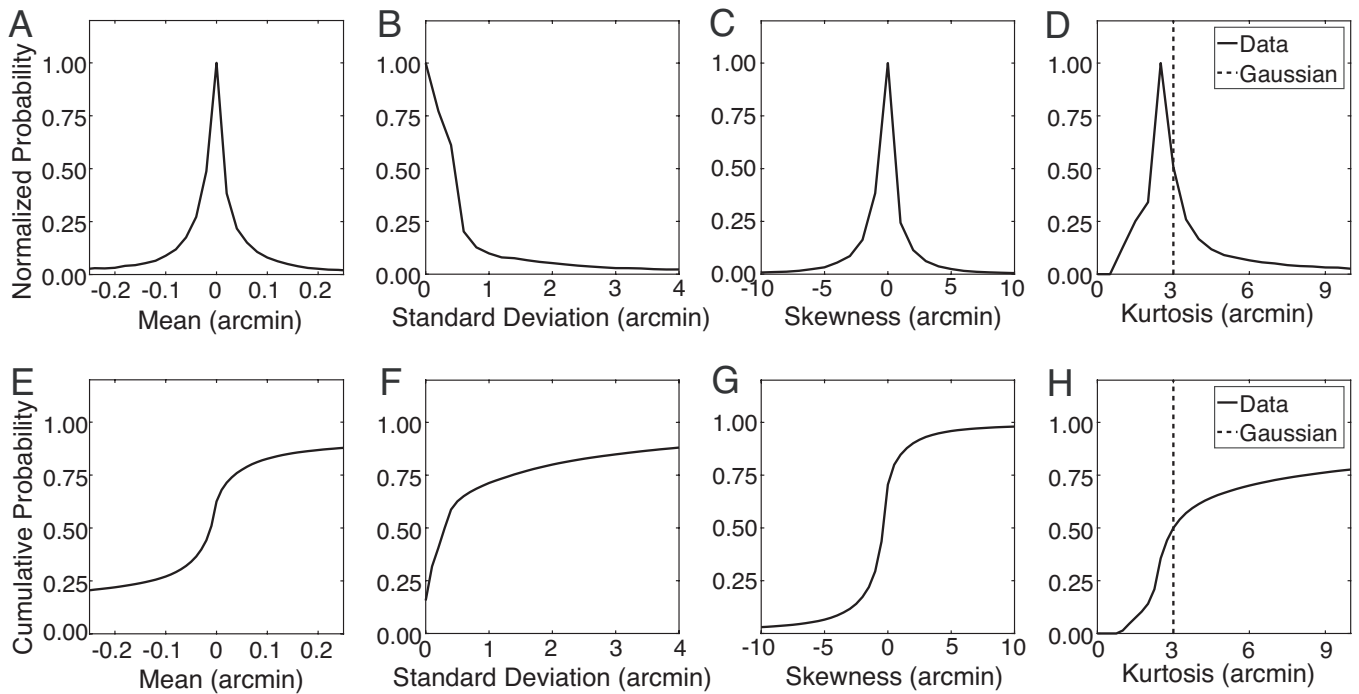


Figure S7. Distributions of within-patch statistics. **A** The distribution of within-patch mean disparity, **B** standard deviation, **C** skewness, and **D** kurtosis, aggregated across all patches in the database. The most likely median and skewness are both zero, suggesting that the within-patch distribution is frequently symmetric about zero. The median skewness is near 0.0 and the median kurtosis is near 3.0, suggesting that within each patch, the median distribution of within-patch disparities is approximately Gaussian. This result suggests that the median within-patch distribution of disparities is well-approximated as Gaussian. However, for a given distribution to be approximately Gaussian, skewness and kurtosis need to take on the appropriate values within the same patch. To check, we computed the joint histogram of within-patch skewness and kurtosis. Results confirm that the median natural image patch is a nearly planar surface with Gaussian distributed disparities.

Supplementary References

- Cooper, E. A., Burge, J., & Banks, M. S. (2011). The vertical horopter is not adaptable, but it may be adaptive. *Journal of Vision*, 11(3). <http://doi.org/10.1167/11.3.20>
- Gibaldi, A., Canessa, A., & Sabatini, S. P. (2017). The Active Side of Stereopsis: Fixation Strategy and Adaptation to Natural Environments. *Nature Publishing Group*, 7, 44800. <http://doi.org/10.1038/srep44800>
- Helmholtz, H. (1925). Treatise on physiological optics. In J. P. C. Shouthall (Ed. & Trans.), *Handbuch der physiologischen optik* (vol. 3). New York: Optical Society of America. (Original work published 1909).
- Schreiber, K. M., Hill, R. L., Filippini, H. R., Schor, C. M., & Banks, M. S. (2008). The surface of the empirical horopter. *Journal of Vision*, 8(3). <http://doi.org/10.1167/8.3.7>
- Sprague, W. W., Cooper, E. A., Tonic, I., & Banks, M. S. (2015). Stereopsis is adaptive for the natural environment. *Science Advances*, 1(4), e1400254–e1400254. <http://doi.org/10.1126/sciadv.1400254>